Reducing the State-Action Complexity in Reinforcement Learning based Traffic Light Control using Self-Organizing Maps

Mladen Miletić, Krešimir Kušić, Martin Gregurić, Edouard Ivanjko

University of Zagreb, Faculty of Transport and Traffic Sciences

www.unizg.hr, http://www.fpz.unizg.hr

mladen.miletic@fpz.hr; kresimir.kusic@fpz.hr; martin.greguric@fpz.hr; edouard.ivanjko@fpz.hr

MOTIVATION FOR ADAPTIVE TRAFFIC LIGHT CONTROL

- **Traffic in urban areas is primarily controlled with traffic light control systems**
  - Fixed traffic signal control
  - Adaptive traffic signal control that changes depending on the current traffic conditions
  - Due to stochastic nature of traffic it is hard to implement optimal traffic light control
  - High state complexity due to many changes in traffic flow directions and volume during the day

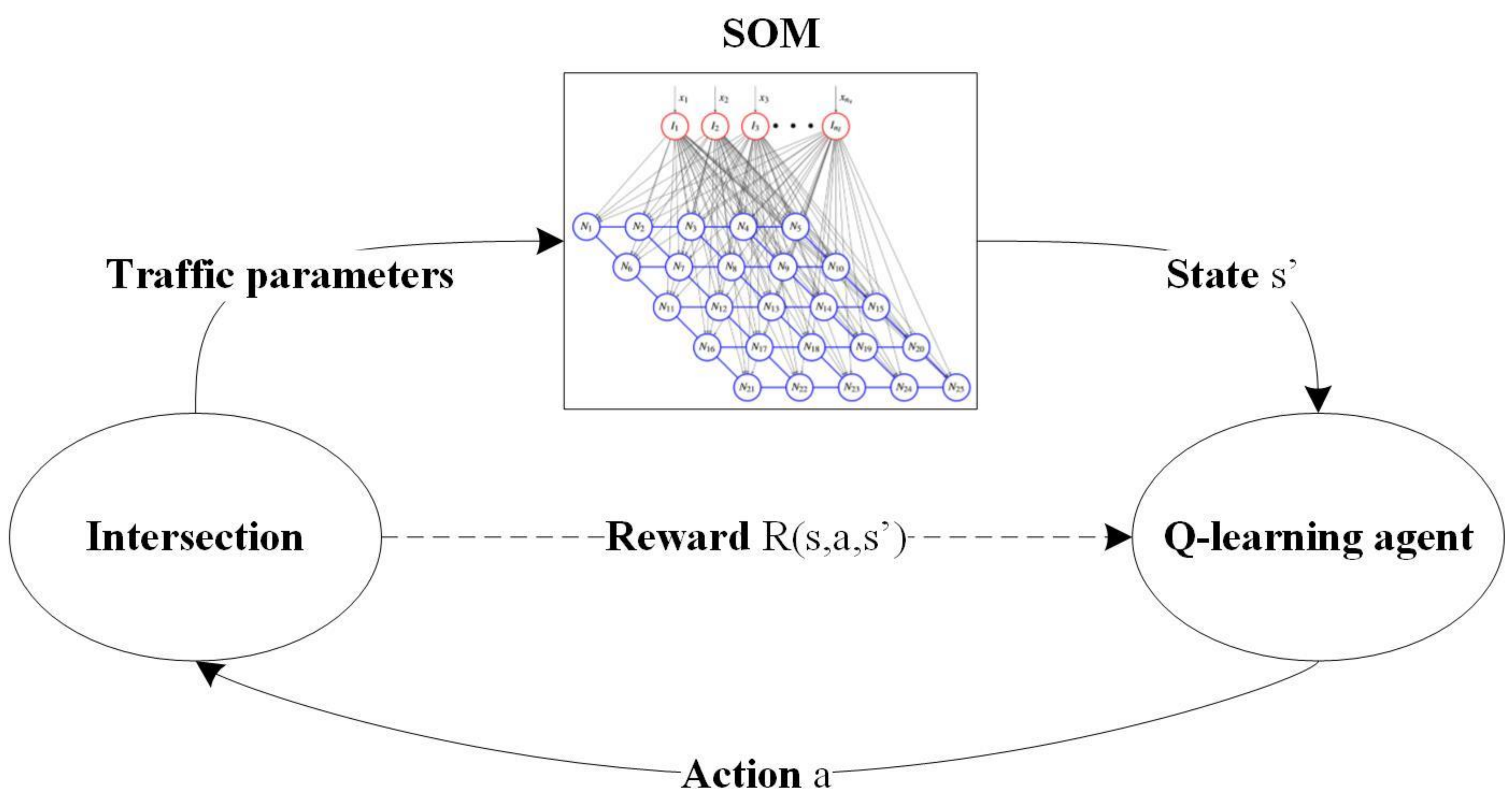
REINFORCEMENT LEARNING USING SELF-ORGANIZING MAPS

- **Self-Organizing Maps (SOM)**
  - A type of neural network with no output layer
  - Weights from input to computational layer represent coordinates in *n*-dimensional space where *n* is the number of neurons in the input layer
  - Unsupervised learning
  - Number of neurons in input layer dependent on the number of measured traffic parameters
  - Number of neurons in computational layer dependent on the desired system resolution (detected traffic states)
  - For a given set of measured traffic parameters, winning neuron is calculated as the neuron with minimal Euclidean distance from the input vector

$$X = (x_1, x_2, x_3, \cdots, x_n) \qquad d(X, W_j) = \sqrt{\sum_{i=1}^{n} (x_i - w_{ij})^2}$$

  - Continuous state space is mapped to discrete state space defined by neuron positions and each neuron encompassing its own neighborhood

- **Reinforcement learning**
  - **Q-learning**
    - **States** (Average and maximum queue lengths for all intersection approaches, processed by SOM)
    - **Actions** (7 distinct signal programs)
    - **Reward** (Change in overall delay of all vehicles)

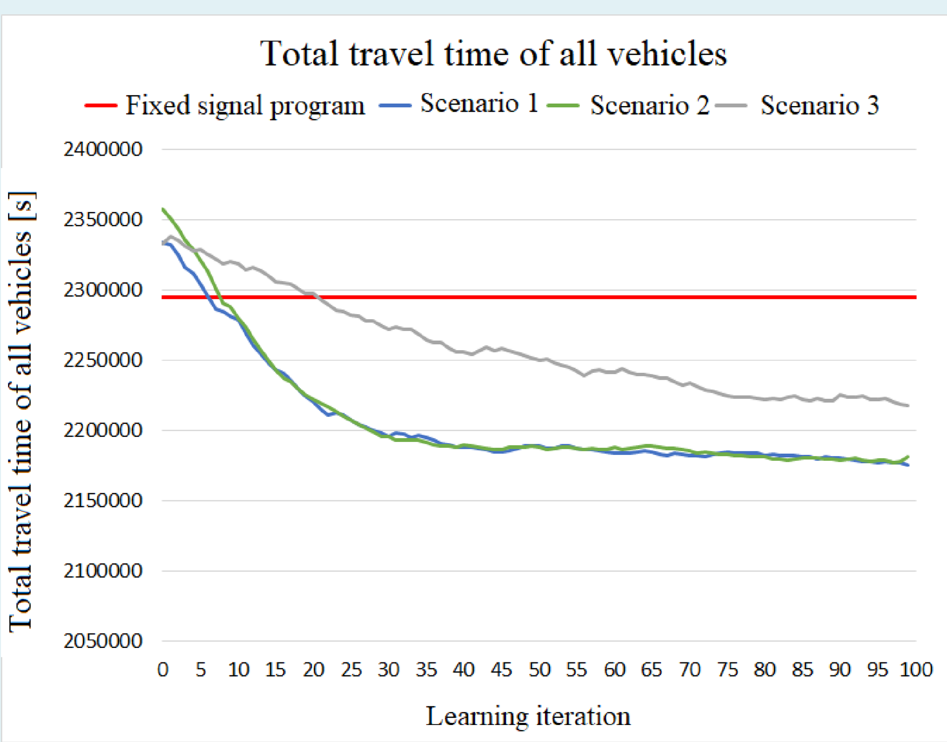


SIMULATION FRAMEWORK AND RESULTS

Evaluation results

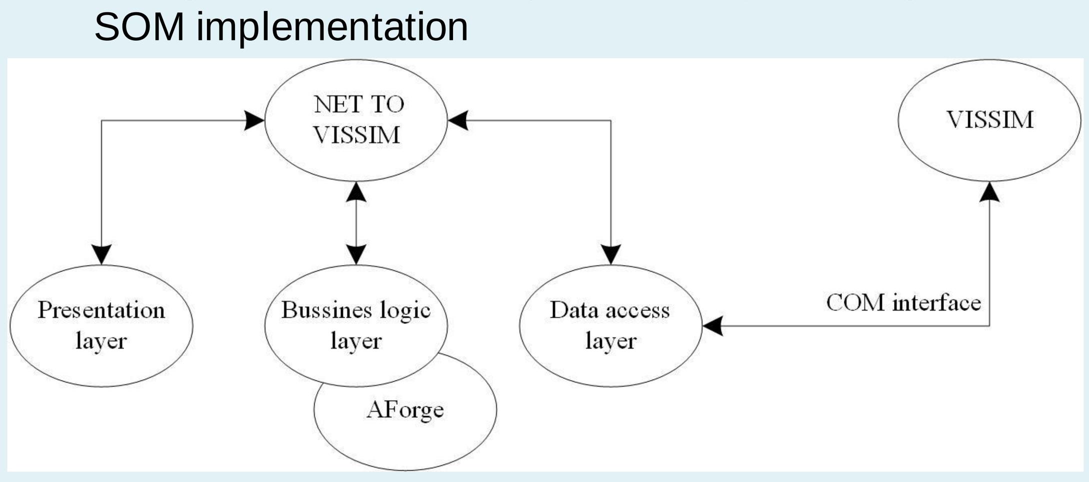
100 simulations, individual duration 15.5 $[h]$)

- **Applied technique**
  - State number reduction using SOM network
  - Q-learning with ε-greedy policy
  - Scenario 1 – 16 neurons – 5.65% reduction
  - Scenario 2 – 9 neurons – 3.61% reduction
  - Scenario 3 – 484 neurons – 3.71% reduction

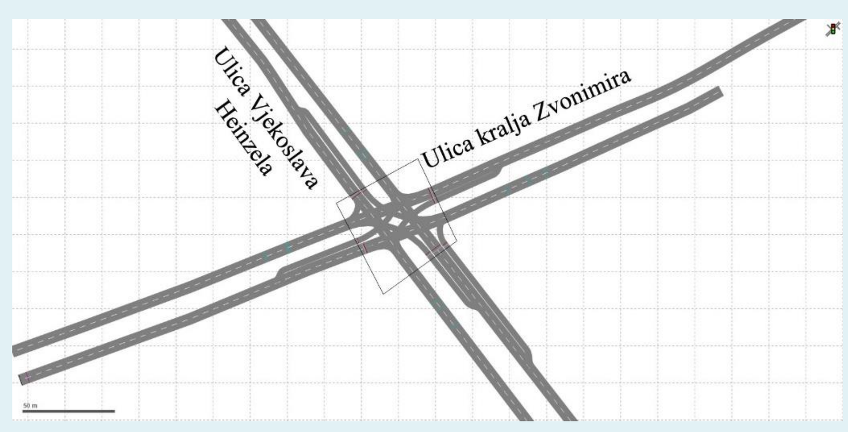


Simulation framework

- VISSIM microscopic simulator used to simulate traffic environment
- NET_TO_VISSIM framework used to control VISSIM objects with COM interface
- AForge framework library used for Q-Learning and SOM implementation



- Isolated intersection simulated using realistic traffic data
- Part of a green wave corridor in the City of Zagreb
- Significant difference in traffic demand of primary and secondary traffic flow
- Simple fixed time signal program operating in three phases



Conclusion

- By implementing SOM the complexity of state space is significantly reduced, which allows reinforcement learning algorithm to converge at a faster rate
- Rate of convergence depends on the number of neurons in SOM
- Total travel time of all vehicles reduced by up to 5.65%
- Average lost time per vehicle reduced by up to 11.54%
- Total number of stops of all vehicles reduced by up to 14.09%
- Total stop time of all vehicles reduced by up to 4.96%

University of Zagreb, Faculty of Transport and Traffic Sciences

Future work on this topic

- Analysis of needed optimal number of neurons for fast convergence and high level of service of controlled intersection
- Implementation for multiple connected intersections
- Analysis of system adaptation to changes in driver/traffic flow behavior
- Cooperation with other Intelligent Transport Systems (ITS) services such as public transport or emergency vehicle preemption

4th International Summer School on Data Science (SSDS 2019), 09th-13th September 2019 Split, Croatia